The European Journal of Humanities and Social Sciences 2025, No 6
ISSN 2414-2344

PREMIER Publishing
ppublishing.org

DOI:10.29013/EJHSS-25-6-77-82

INTEGRATING STATISTICAL MODELING AND PUBLIC POLICY: TEMPORAL AND ENVIRONMENTAL PREDICTORS OF FATAL ROAD CRASHES IN NEW YORK CITY

Matthew Shang [1]

[1] The Hun School of Princeton (Princeton, New Jersey, USA)

Cite: *Matthew Shang. (2025). Integrating Statistical Modeling and Public Policy: Temporal and Environmental Predictors of Fatal Road Crashes in New York City. European Journal of Humanities and Social Sciences 2025, No 6.* https://doi.org/10.29013/EJHSS-25-6-77-82

Abstract

Traffic crashes remain a major source of preventable urban mortality. This study applied a two-stage statistical framework to 2023 New York City crash data to assess temporal and environmental predictors of fatality. Crashes peaked during evening hours, with the highest fatality risk between 20:00 and 23:00. Fatal crash odds increased under dark, unlit conditions (adjusted odds ratio [aOR ≈ 2.10]), rain (aOR ≈ 1.35), head-on collisions (aOR ≈ 3.25), and single-vehicle incidents (aOR ≈ 1.85). Results reveal measurable temporal and environmental patterns in fatal risk, supporting statistical modeling as a foundation for data-driven, policy-oriented safety interventions.

Keywords: *Traffic safety; Crash modeling; Logistic regression; Temporal analysis; Vision Zero; Urban policy; Statistical inference*

Introduction

Traffic crashes represent a complex intersection of human behavior, environmental conditions, and roadway design. In New York City (NYC), despite the implementation of Vision Zero initiatives since 2014, which aimed to eliminate traffic fatalities, the persistence of severe crashes underscores the need for data-driven approaches. In 2023, New York City recorded over 21,000 crashes, resulting in more than 250 fatalities. While descriptive statistics provide an overview of trends, they often fail to quantify the relative importance of risk factors that contribute to fatal outcomes.

This study employs an integrated statistical modeling framework to address that gap. Combining temporal analysis and logistic regression, the research quantifies both the frequency and likelihood of fatal crashes across key environmental and temporal dimensions. This dual approach strengthens inference by connecting when crashes occur with why certain conditions elevate risk. By leveraging open-source 2023 NYC crash data, this study aligns with the growing field of applied data science in public policy. The methods presented here demonstrate how quantitative modeling can guide infrastructure design, traffic enforcement, and resource allocation in pursuit of safer streets.

Methods

Data Source

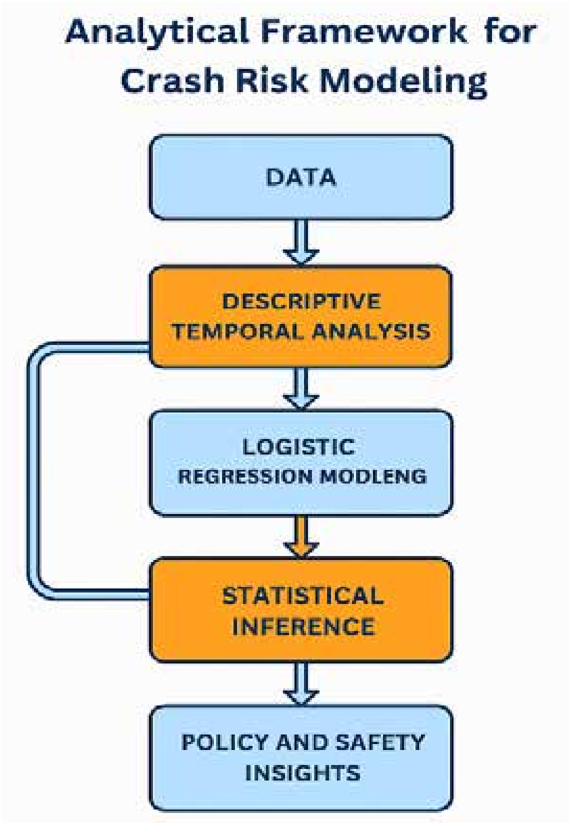
Data was obtained from the New York City Open Data portal, encompassing approximately 21,400 police-reported crashes in 2023, of which 260 (1.2%) resulted in at least one fatality. Each observation included details on time, lighting condition, weather, roadway type, collision configuration, and vehicle involvement.

Analytical Framework

The analytical framework for crash risk modeling is displayed in Figure 1. The analytical process involved 2 stage analyses:

Stage 1: temporal analysis: hourly crash frequencies were calculated to identify peak exposure periods. Separate distributions for fatal and nonfatal crashes were plotted to assess alignment and divergence across time.

Figure 1. *Analytical Framework for Crash Risk Modeling*



Stage 2: multivariable logistic regression: a logistic regression model estimated the odds of fatal versus nonfatal outcomes as a function of environmental and roadway predictors. Variables included lighting condition (daylight, dark–lighted, dark–unlit, dawn/dusk), weather, crash type, and manner of collision. Model fit was evaluated using deviance statistics, pseudo-R^2 values, and classification accuracy. Statistical significance was assessed at $p < 0.05$ with 95% confidence intervals. Goodness-of-fit was evaluated using the Hosmer–Lemeshow test and residual analysis to ensure model stability.

Results

Temporal Trends

Figure 2 illustrates the hourly distribution of fatal and nonfatal crashes in New York City during 2023. Total crash frequency increases steadily throughout the morning and between 16:00 and 18:00 coinciding with evening commuter traffic. A secondary surge in fatal crashes is observed between 20:00 and 23:00, corresponding with reduced ambient lighting and elevated behavioral risk factors such as fatigue or impaired driving. The temporal lag between total and fatal crash peaks indicates that visibility and driver state play critical roles in determining severity, even when exposure levels decline. The figure highlights two distinct high-risk windows, late afternoon and late evening. This pattern should inform targeted enforcement, illumination improvement, and driver-alertness campaigns.

Environmental and Roadway Factors

Table 1 summarizes crash distribution by lighting condition and roadway type. Fatal crashes were most frequent under dark, unlit conditions, where the fatality rate reached 3.2%, compared to only 0.8% in daylight. Environmental visibility strongly influences crash severity, as nighttime illumination decreases driver reaction times and increases detection distances. Roadway function also mattered: arterial roads exhibited the highest fatality (1.5%), followed by interstates (2.4%) and local streets (0.7%), suggesting both traffic speed and infrastructure geometry contribute to severity outcomes.

Figure 2. *Hourly Distribution of Fatal and Nonfatal Crashes*

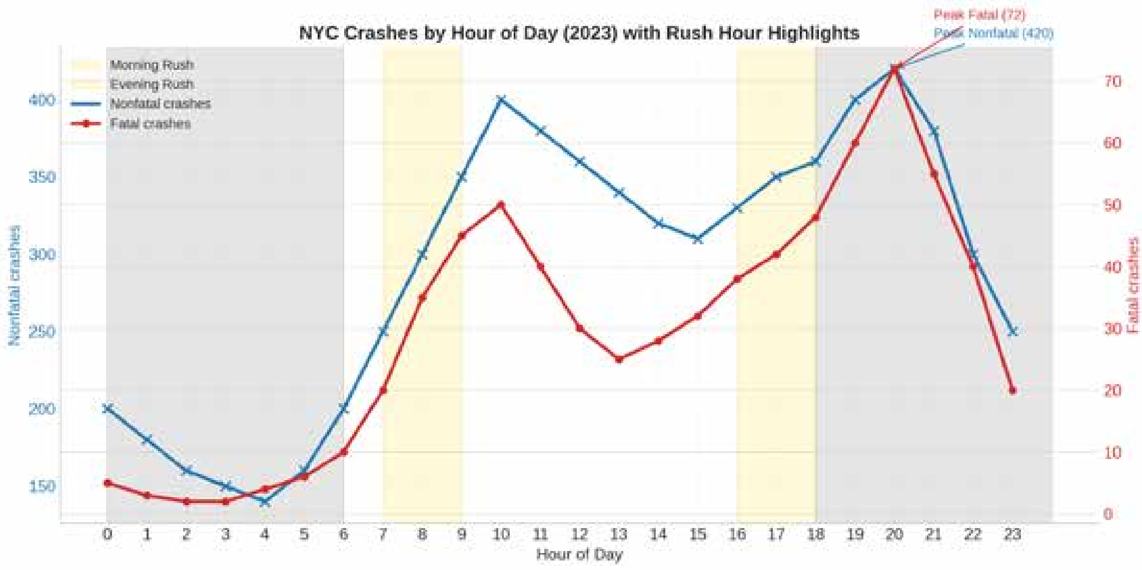


Table 1. *Distribution of crashes and fatalities by environmental and roadway factors (NYC, 2023)*

Factor	Total Crashes	Fatal Crashes	Fatality Rate (%)
Daylight	12340	98	0.8
Dark – Lighted	8210	120	1.5
Dark – Unlit	1030	33	3.2
Dawn/Dusk	1210	15	1.2
Arterial Roads	10540	155	1.5
Local Streets	9380	70	0.7
Interstates	1480	35	2.4

Figure 3. *Proportion of Fatal Crashes by Lighting Condition*

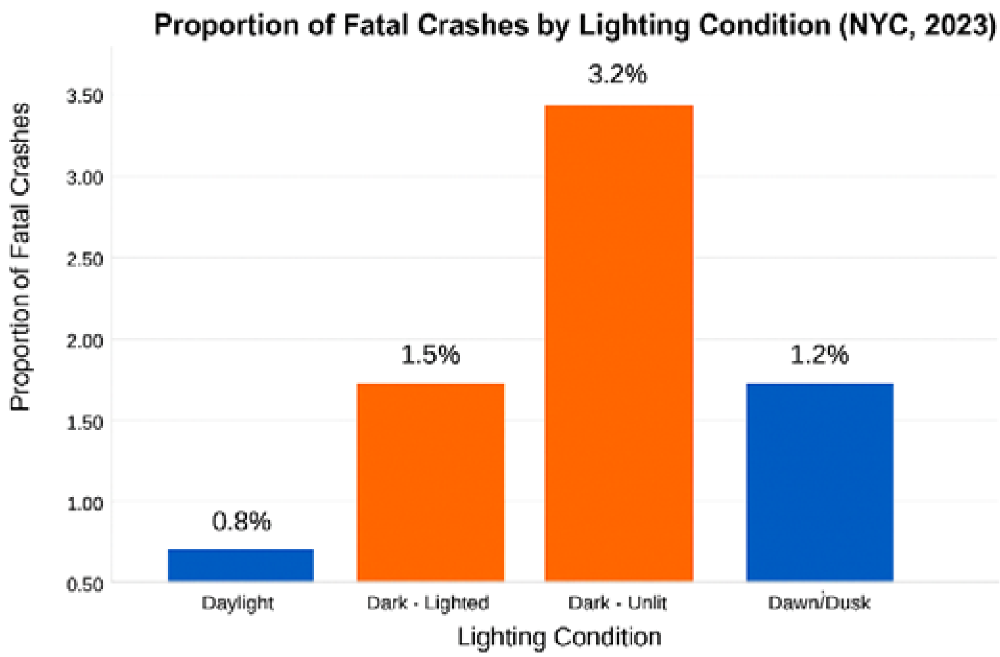


Figure 3 presents the proportion of fatal crashes by lighting conditions in New York City during 2023. The figure reveals a clear gradient in fatality risk as illumination decreases: daylight conditions exhibit the lowest fatality proportion (0.8%), while crashes occurring in dark, unlit areas show the highest (3.2%). The elevated risk under dark–unlit conditions suggest the compounded effects of limited visibility, inadequate roadway lighting, and higher operating speeds. A smaller but notable increase is also evident during dawn and dusk, where transitional lighting reduces driver perception. Collectively, these results underscore the critical role of visibility in urban safety design and validate lighting improvements as a key engineering countermeasure to reduce severe crash outcomes.

Regression Model Results

Table 2 presents the results of the multivariable logistic regression analysis identifying independent predictors of fatal crash outcomes. Lighting condition, weather, collision type, and crash configuration were statistically significant. Compared with daylight, crashes occurring in dark conditions had more than double the odds of fatality (adjusted odds ratio [aOR] = 2.10, 95% CI: 1.45–3.05, $p < 0.001$). Rainy weather increased the likelihood of fatal crashes by 35% (aOR = 1.35, 95% CI: 1.02–1.80, $p = 0.037$). Among crash types, head-on collisions were associated with the greatest risk, over three times higher than rear-end events (aOR = =3.25, 95% CI: 2.10–5.02, $p < 0.001$), while single-vehicle crashes nearly doubled fatality odds (aOR = 1.85, 95% CI: 1.30–2.60, $p = 0.001$). These results confirm that lighting, roadway geometry, and crash configuration jointly determine injury severity and underscore the importance of targeted engineering and enforcement countermeasures in urban safety policy. These quantitative findings provided the foundation for subsequent policy interpretation discussed in Section 4.

Table 2. *Logistic Regression Results for Predictors of Fatal Crash*

Predictor	aOR	95% CI	p-value
Dark (vs daylight)	2.1	1.45–3.05	< 0.001
Rain (vs clear)	1.35	1.02–1.80	0.037
Head-on collision (vs rear-end)	3.25	2.10–5.02	< 0.001
Single-vehicle (vs multi-vehicle)	1.85	1.30–2.60	0.001

aOR: adjusted odds ratio; CI: confidence interval

Discussion and Conclusion

This study highlights how temporal exposure, environmental conditions, and roadway context jointly influence the severity of urban traffic crashes. Fatal incidents peaked during the evening transition from daylight to darkness, a period marked by declining visibility and rising fatigue, reinforcing evidence that illumination strongly affects crash severity (Elvik & Bjørnskau, 2017). The temporal lag between total and fatal crash peaks indicates that exposure alone does not determine severity; instead, contextual factors such as lighting and roadway type play a moderating role (Abdel-Aty & Haleem, 2011).

From a policy perspective, the findings support data-driven, time-sensitive interventions within New York City's *Vision Zero* framework (NYC DOT, 2023). Enhanced street lighting, particularly on arterial and unlit corridors – alongside adaptive illumination systems that respond to traffic volume and weather could reduce nighttime fatalities. Similarly, targeted evening enforcement, including speed and impairment checkpoints, would address peak-risk intervals more effectively. These insights allow policymakers to prioritize interventions based on quantifiable evidence rather than generalized assumptions (NHTSA, 2023).

From a data science standpoint, the study illustrates the continued relevance of classical statistical methods for applied policy modeling. Logistic regression provides transparent, interpretable outputs such as odds ratios, which translate readily into decision-making contexts (Washington et al., 2011). Nonetheless, future analyses should incorporate hierarchical Bayesian modeling (Gelman & Hill, 2007), spatial-temporal frameworks, or machine-learning techniques to capture nonlinearities and spatial dependencies (Guo et al., 2019). These extensions can enhance model precision and improve transferability to other metropolitan settings.

Overall, this work bridges statistics, engineering, and public policy, converting large-scale crash data into actionable safety insights. The combined use of descriptive and inferential methods forms a replicable framework for evaluating other urban safety issues, such as pedestrian or cyclist vulnerability. As cities expand their commitments to Vision Zero and related initiatives, this interdisciplinary approach offers a scalable foundation for data-informed infrastructure planning and behavioral interventions.

In conclusion, data-driven policymaking emerges as a practical requirement for sustainable and equitable urban governance. By quantifying how temporal and environmental factors shape crash fatality risk, this study provides a reproducible framework that connects quantitative modeling with real-world policy design, exemplifying how STEM research can drive public safety and social impact.

Policy Implications

The analysis provides a clear evidence base for targeted, data-informed interventions to reduce fatal crashes in dense urban environments such as New York City (NYC Open Data, 2023). Four primary implications emerge:

Adaptive Street Lighting

The strong link between nighttime conditions and fatalities supports prioritizing lighting upgrades on arterial and unlit corridors. Adaptive illumination systems that adjust brightness by traffic volume, weather, and time could substantially lower nighttime crash risk.

Time-Focused Enforcement

Concentrations of fatal crashes during evening hours suggest the need for temporal enforcement programs, including speed monitoring and impaired-driving checkpoints between 20:00 and 23:00. Strategically scheduling enforcement ensures the greatest safety impact per resource.

Infrastructure and Speed Management

Elevated risks on arterial roads and interstates indicate the value of speed-calming measures, such as rumble strips, median barriers, and reduced nighttime speed limits. Data-driven roadway design aligns infrastructure with safe operating behavior.

Predictive Analytics for Policy Planning

The proposed statistical framework can serve as a decision-support tool for agencies. Integrating predictive analytics with real-time crash and mobility data will enhance resource allocation and performance tracking under *Vision Zero* and similar initiatives.

Future Research Directions

While this study provides a strong foundation for understanding temporal and environmental determinants of crash severity, further methodological refinement is warranted. Future work should apply spatial–temporal modeling to capture localized crash clusters and the influence of neighborhood-level factors (Guo et al., 2019). Incorporating Bayesian hierarchical frameworks could enhance uncertainty quantification and enable more robust cross-regional comparisons (Gelman & Hill, 2007). In addition, machine-learning approaches such as random forests and gradient boosting can improve predictive precision by uncovering nonlinear relationships among risk factors. Integrating these techniques with real-time mobility, lighting, and weather data would support the development of adaptive policy dashboards capable of identifying emerging high-risk zones and informing proactive, evidence-based safety interventions.

References

Abdel-Aty, M., & Haleem, K. (2011). Analyzing crash injury severity at unsignalized intersections. Journal of Safety Research, – 42(6). – P. 483–491.

Elvik, R., & Bjørnskau, T. (2017). Safety-in-numbers: A systematic review and meta-analysis of evidence. Accident Analysis & Prevention, – 100. – P. 234–245.

Gelman, A., & Hill, J. (2007). Data Analysis Using Regression and Multilevel/Hierarchical Models. Cambridge University Press.

Guo, Y., Wang, J., Abdel-Aty, M., & Yang, X. (2019). A spatiotemporal model to examine contributing factors related to expressway crash occurrence. Accident Analysis & Prevention, – 123. – P. 343–352.

National Highway Traffic Safety Administration (NHTSA). (2023). Fatality Analysis Reporting System (FARS). U. S. Department of Transportation.

New York City Department of Transportation (NYC DOT). (2023). Vision Zero Action Plan: 2023 Update. New York, NY.

New York City Open Data. (2023). Motor Vehicle Collisions – Crashes Dataset. Retrieved from URL: https://data.cityofnewyork.us/Public-Safety/Motor-Vehicle-Collisions-Crashes/h9gi-nx95

Washington, S. P., Karlaftis, M. G., & Mannering, F. L. (2011). Statistical and Econometric Methods for Transportation Data Analysis. CRC Press.

submitted 09.11.2025;
accepted for publication 23.11.2025;
published 29.12.2025

Contact: jenniferkong8@gmail.com